Confocal Scanning Laser Ophthalmoscopy Classifiers and Stereophotograph Evaluation for Prediction of Visual Field Abnormalities in Glaucoma-Suspect Eyes

Christopher Bowd,[1] *Linda M. Zangwill,*[1] *Felipe A. Medeiros,*[1] *Jiucang Hao,*[2,3] *Kwokleung Chan,*[2,3] *Te-Won Lee,*[2,3] *Terrence J. Sejnowski,*[2,3] *Michael H. Goldbaum,*[1] *Pamela A. Sample,*[1] *Jonathan G. Crowston,*[1] *and Robert N. Weinreb*[1]

Purpose. To determine whether Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Dossenheim, Germany) classification techniques and investigational support vector machine (SVM) analyses can detect optic disc abnormalities in glaucoma-suspect eyes before the development of visual field abnormalities.

Methods. Glaucoma-suspect eyes ($n = 226$) were classified as converts or nonconverts based on the development of repeatable (either two or three consecutive) standard automated perimetry (SAP)–detected abnormalities over the course of the study (mean follow-up, ~4.5 years). Hazard ratios for development of SAP abnormalities were calculated based on baseline classification results, follow-up time, and end point status (convert, nonconvert). Classification techniques applied were HRT classification (HRTC), Moorfields Regression Analysis, forward-selection optimized SVM (SVM fwd) and backward elimination–optimized SVM (SVM back) analysis of HRT data, and stereophotograph assessment.

Results. Univariate analyses indicated that all classification techniques were predictors of the development of two repeatable abnormal SAP results, with hazards ratios (95% confidence interval [CI]) ranging from 1.32 (1.00–1.75) for HRTC to 2.0 (1.48–2.76) for stereophotograph assessment (all $P \leq 0.05$). Only SVM (SVM fwd and SVM back) analysis of HRT data and stereophotograph assessment were univariate predictors of the development of three repeatable abnormal SAP results, with hazard ratios (95% CI) ranging from 1.73 (1.16–2.82) for SVM fwd to 1.82 (1.19–3.12) for SVM back (both $P < 0.007$). Multivariate analyses including each classification technique individually in a model with age, baseline SAP pattern standard deviation [PSD], and baseline IOP indicated that all classification techniques except HRTC ($P = 0.06$) were predictors of the development of two repeatable abnormal SAP results with hazards ratios ranging from 1.30 (0.99, 1.73) for HRTC to 1.90 (1.37, 2.69) for stereophotograph assessment. Only SVM (SVM fwd and SVM back) analysis of HRT data and stereophotograph assessment were significant predictors of the development of three repeatable abnormal SAP results in multivariate analyses; hazard ratios of 1.57 (1.03, 2.59) and 1.70 (1.18, 2.51), respectively. SAP PSD was a significant predictor of two repeatable abnormal SAP results in multivariate models with all classification techniques, with hazard ratios ranging from 3.31 (1.39, 7.89) to 4.70 (2.02, 10.93) per 1-dB increase.

Conclusions. HRT classifications techniques and stereophotograph assessment can detect optic disc topography abnormalities in glaucoma-suspect eyes before the development of SAP abnormalities. These data support strongly the importance of optic disc examination for early glaucoma diagnosis. (*Invest Ophthalmol Vis Sci.* 2004;45:2255–2262) DOI:10.1167/iovs.03-1087

Optic disc damage has been shown to precede visual field loss measured using standard automated perimetry (SAP).[1–3] To detect optic disc topography changes that occur before detectable visual field loss by using optical imaging and other techniques, longitudinal follow-up of patients with glaucoma is necessary. Although investigators in several studies have examined the cross-sectional effect of age on optic disc topography using confocal scanning laser ophthalmoscopy (CSLO),[4–8] few have examined the ability of CSLO to detect change in optic disc topography over time. One study showed concurrent CSLO parameter changes and SAP changes in ocular hypertensive eyes in which visual field defects developed.[9] Another study showed significant changes in rim area before conversion in ocular hypertensive eyes in which visual field defects developed.[10] Finally, a third study showed progressive change of retinal height in glaucomatous eyes that also showed change on expert stereophotograph evaluation and SAP.[11]

In the present study, the ability of CSLO to detect optic disc topography abnormalities in glaucoma-suspect eyes before the development of repeatable visual field changes was examined. Because changes in optic disc topography may precede SAP-detectable visual field defects, we hypothesized that suspect eyes classified at baseline as having optic discs outside of normal limits would be more likely than those classified as within normal limits to undergo development of visual field defects. In addition, we sought to determine whether results from newly applied machine learning classifiers were better predictors of glaucomatous visual field abnormalities than currently available HRT software and stereophotograph evaluation.

Methods

Subjects

A randomly selected eye from each of 226 glaucoma-suspect patients enrolled in the longitudinal Diagnostic Innovations in Glaucoma Study

From the [1]Hamilton Glaucoma Center of the Department of Ophthalmology, and the [2]Institute for Neural Computation, University of California San Diego, La Jolla, California; and the [3]Computational Neurobiology Laboratory, The Salk Institute, La Jolla, California.

Supported by The Glaucoma Research Foundation (CB); National Eye Institute Grants EY11008 (LMZ), EY13235 (MHG), EY08208 (PAS), and EY010401 (RNW); and the Foundation for Eye Research (FAM).

Submitted for publication October 1, 2003; revised March 3, 2004; accepted March 23, 2004.

Disclosure: **C. Bowd**, None; **L.M. Zangwill**, None; **F.A. Medeiros**, None; **J. Hao**, None; **K. Chan**, None; **T.-W. Lee**, None; **T.J. Sejnowski**, None; **M.H. Goldbaum**, None; **P.A. Sample**, Carl Zeiss Meditec (F); **J.G. Crowston**, None; **R.N. Weinreb**, Heidelberg Engineering (F), Carl Zeiss Meditec (F)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Christopher Bowd, Hamilton Glaucoma Center, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0946; cbowd@eyecenter.ucsd.edu.

(DIGS) at the Hamilton Glaucoma Center of the University of California, San Diego, was studied. Glaucoma-suspect patients had a repeatable intraocular pressure of more than 22 mm Hg and/or glaucomatous-appearing optic discs on examination. Selected eyes were those with a minimum of four full-threshold or Swedish Interactive Threshold Algorithm (SITA) SAP tests, an initial normal SAP result at the time of study enrollment, and at least 2 years of HRT examination follow-up. Ninety-one percent (206/226) of eyes had at least two consecutive normal SAP results on the tests closest to HRT baseline. Of the 20 that did not, 15 (7% of total eyes) had two of three normal SAP results (i.e., a normal, abnormal, normal sequence), and five (2% of total eyes) had one normal SAP result followed by two or more abnormal ones.

These eyes were classified into convert and nonconvert groups based on the development of repeatable SAP abnormalities over the course of the study based on two different criteria. The first criterion required convert eyes to exhibit two consecutive abnormal SAP results (called convert 2 criterion): defined as a glaucoma hemifield test outside of normal limits or a pattern standard deviation $\geq$ 95% of the normal population (global indices were used, in part, because both SAP full-threshold and SAP SITA test strategies were used to determine conversion). Nonconvert eyes did not show two consecutive abnormal SAP results. The second criterion required convert eyes to show three consecutive abnormal SAP results (convert 3 criterion). Nonconvert eyes did not show three consecutive abnormal SAP results. These two criteria were used to make our results comparable to those previously reported, because some studies have required two consecutive abnormal results to classify eyes as glaucomatous,[12,13] and the recent Ocular Hypertension Treatment Study required three consecutive abnormal results to define conversion to glaucoma.[14] The five eyes that had a single normal baseline result followed by two or more abnormal ones were classified as converts because, in a clinical setting, a single normal baseline visual field test result is most often not confirmed for diagnostic purposes.

For both criteria, eyes underwent a baseline HRT examination at least 1 year before the end point. The end point for converts was defined as the first of two (convert 2 criterion) or three (convert 3 criterion) consecutive abnormal visual field test results, and the end point for nonconverts was defined as the last available visual field result. Informed consent was obtained from all participants, and all methods were approved by the University of California, San Diego, Institutional Review Board (IRB) and adhered to the guidelines of Declaration of Helsinki for research involving human subjects.

Before HRT imaging, all participants underwent a complete ophthalmic examination, including slit lamp biomicroscopy, IOP measurement by applanation tonometry, and dilated stereoscopic ophthalmoscopy with a hand-held 78-D lens. Only eyes with visual acuity of 20/40 or better and refractive error within $\pm$5.00 D were included. Eyes with coexisting retinal disease, uveitis, or nonglaucomatous optic neuropathy were excluded. All participants were 40 years of age or older at the time of baseline HRT imaging. The average age of all the participants enrolled in the study was 61.7 $\pm$ 10.7 years (SD; range, 40.0–88.1). Fifty-nine percent of participants were women and 89% were white.

Confocal Scanning Laser Ophthalmoscopy

The Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Dossenheim, Germany) employs confocal scanning diode technology to provide topographical measures of the optic disc and parapapillary retina. This instrument has been described in detail elsewhere.[15]

Three 15° field-of-view scans, centered on the optic disc and judged to be of acceptable quality by a trained technician were obtained for each tested eye. A mean topography image of these three scans was created using HRT software version 2.01. A trained technician outlined the optic disc margin on the mean image using information obtained by viewing simultaneous stereoscopic photographs of the optic disc. All mean topography images had a standard deviation less than 50 μm.

HRT Classification

We determined the value of one investigational and two currently available HRT classification techniques for predicting the development of glaucomatous visual field abnormalities based on baseline measurements. Eyes were classified as either abnormal or normal at baseline according to the following.

HRT Classification. The HRTC classification (HRTC) is a software-provided linear discriminant function including the parameters rim volume, height variation contour, and cup shape (adjusted for age).[16] According to HRT software, eyes with negative values were classified as abnormal, and eyes with positive values were classified as normal.

Moorfields Regression Analysis Classification. The Moorfields Regression Analysis (MRA) compares the log of the measured neuroretinal rim area (globally and regionally) with the predicted age-corrected rim area (adjusted for optic disc area) derived from a normative database.[17] MRA was evaluated globally and regionally. Regions were defined as temporal superior (46–90°), nasal superior (91–135°), nasal (136–225°), nasal inferior (226–270°), temporal inferior (271–315°), and temporal (316–45°). If the measured log rim area was lower than the software-specified 99.9% cutoff, based on healthy eyes, for the entire disc or for any of the regions, the eye was classified as abnormal.

Gaussian Support Vector Machine Classification. The support vector machines (SVMs) are machine classification techniques used for solving classification and regression problems. These techniques have been applied to HRT data and have been shown to discriminate between glaucomatous (by visual field abnormality) and healthy eyes as well as, or better than, linear discriminant function techniques and other machine-learning classifier techniques.[12] Specifics of this technique have been described elsewhere.[18,19] SVM cutoff points used to classify eyes in the present study were derived from a previous study[20] that investigated the ability of SVM to discriminate between 135 healthy eyes (on examination and by SAP) and 95 glaucomatous eyes (those with two consecutive SAP abnormalities).

In the previous study, SVM was trained on a set of HRT topographic parameters from labeled eyes (healthy versus glaucoma) that included the following global measurements: disc area, area below reference, mean height of contour, peak height of contour, height variability of contour, volume below surface, volume above surface, volume below reference, volume above reference, maximum cup depth, cup shape, mean cup depth, retinal nerve fiber layer (RNFL) thickness, RNFL cross-sectional area, reference height, rim area, cup-to-disc ratio, and rim-to-disc ratio. Also included were six regional topographic measures of each parameter listed above except mean cup depth, RNFL cross-sectional area, and reference height. Regions were the same as those used for MRA in the present study. All these parameters have been discussed in more detail elsewhere.[21] In addition, measurements of the mean height of the operator-drawn contour line surrounding the optic disc margin (mean height contour at disc margin) obtained in thirty-six 10° sectors (with sector 1 corresponding with the horizontal temporal position, 0–9° unit circle) and measurements of the mean retinal height at a location in the parapapillary region 1.7 times the disc margin radius in thirty-six 10° sectors were used. The sectoral mean height contour at disc margin measurements describe the height of the neuroretinal rim at the optic disc margin, and the sectoral mean retinal height measurements describe the retinal height in the parapapillary region, independent of the standard reference plane.

Next, in the previous study, forward selection and backward elimination were used to determine two subsets of all topographic parameters (full dimensional data set) that resulted in the peak receiver operator characteristic (ROC) curve areas (ROC curves were not always maximized using all the parameters in the full dimensional data set). Optimized SVMs using both forward selection and backward elimination showed similar classification performance (ROC curve areas similar), so both optimized SVM classifiers were applied to the separate data set used in the present study.

TABLE 1. Demographic and Ocular Characteristics

	Convert 2			Convert 3		
	Converts (n = 51)	Nonconverts (n = 175)	P	Converts (n = 37)	Nonconverts (n = 189)	P
Age (y, SD)	64.1 (12.0)	61.0 (10.2)	0.070	63.2 (12.3)	61.4 (10.3)	0.35
Percent white	82	91	0.120	84	90	0.26
Percent female	57	60	0.750	59	59	0.98
Average follow-up (y, SD)*	4.2 (2.0)	4.7 (2.6)	0.140	4.6 (1.9)	4.9 (2.7)	0.49
Closest SAP PSD (dB, SD)	1.92 (0.31)	1.73 (0.33)	<0.001	1.87 (0.31)	1.75 (0.34)	0.04
Closest IOP (mm Hg, SD)	23.3 (5.0)	23.3 (5.6)	0.970	23.4 (5.0)	23.3 (5.6)	0.87
Participants with IOP-lowering drugs or procedure prior to baseline HRT (%)	18 (35%)	53 (30%)	0.500	12 (32%)	59 (31%)	0.99
Participants with IOP-lowering drugs or procedure during follow-up (%)	33 (65%)	97 (55%)	0.260	24 (65%)	106 (56%)	0.37

* Follow-up time is years from baseline to the first of two or three consecutive abnormal field test results for convert eyes and years from baseline to final study field in nonconvert eyes.

To classify eyes as abnormal with SVMs in the present study, we used two cutoffs with specificities of 95% and 85% established from the ROC curve results of the previous study. We chose these specificity values because, in the 135 healthy eyes from the previous study, specificities of HRTC, MRA, and stereophotography were 86%, 93%, and 84%, respectively. Ninety-five percent and 85% were chosen to approximate these values for the best comparison of the baseline predictive values of SVM with those of HRTC, MRA, and stereophotograph classification techniques for identifying convert eyes.

Eyes were considered abnormal in the present study if the SVM output value exceeded the cutoff used for the forward-selection data subset or for the backward elimination data subset from the previous study. Therefore, results from four SVM classification schemes applied to the current data are reported. These include:

1. The forward-selection–identified data subset using the 85% specificity cutoff (SVM fwd 85)
2. The forward-selection–identified data subset using the 95% specificity cutoff (SVM fwd 95)
3. The backward-elimination–identified data subset using the 85% specificity cutoff (SVM back 85)
4. The backward-elimination–identified data subset using the 95% specificity cutoff (SVM back 95).

All SVM analyses were conducted on computer (MatLab ver. 5.0; The MathWorks, Inc., Natick, MA).

Stereophotography Classification

We also determined the value of masked expert assessment of simultaneous stereoscopic optic disc photographs for predicting the development of glaucomatous visual field abnormalities. Photographs were obtained at the time of baseline HRT measurements. A fundus camera (TRC-SS camera; Topcon Instrument Corp. of America, Paramus, NJ) was used after maximum pupil dilation, and all photograph evaluations were performed with a stereoscopic viewer (Pentax II; Asahi Optical Co., Tokyo, Japan) used by experienced graders who assessed each photograph independently in a masked fashion. All eyes were graded subjectively as either abnormal (i.e., glaucomatous optic neuropathy) or normal based on the presence of rim thinning, focal or diffuse RNFL thinning, or nonphysiologic excavation and/or undermining of the cup, or cup-disc ratio asymmetry of more than 0.2. If the two graders did not agree on the classification of an eye, another independent grade was solicited to reach a majority decision. Although optic disc damage is probably a continuous variable (as are the other classification techniques used), we used a dichotomous (abnormal versus normal) decision, because HRTC and MRA provide dichotomous output.

Standard Automated Perimetry

Study eyes were tested annually (Humphrey Visual Field Analyzer; Carl Zeiss Meditec, Dublin, CA) using the Humphrey 24-2 or 30-2 full-threshold or SITA strategies. Tests were considered reliable if all reliability indices were less than 33%.

Analyses

Cox proportional hazards regression (using JMP software, SAS Institute, Inc., Cary, NC) was used to calculate the univariate hazard ratios between conversion to SAP abnormality (for both convert 2 and convert 3 criteria) and baseline classification based on HRTC, MRA, SVM fwd 85, SVM fwd 95, SVM back 85, SVM back 95, and stereophotograph assessment. We also assessed the multivariate association between each classification technique and conversion to visual field abnormality with other possible predictors, including age at baseline HRT date, closest recorded SAP PSD, and closest IOP measurement to baseline HRT date in the model. Closest IOP was selected to control for the effect of IOP at the time of HRT imaging, because HRT topography can be affected by changes in IOP.[22–24] All these variables were significant independent predictors of the development of primary open angle glaucoma in the Ocular Hypertensive Treatment Study.[14]

RESULTS

Convert 2 Criterion

Fifty-one (23%) of 226 eyes had two consecutive abnormal SAP results, and 175 (77%) of 226 eyes did not. Demographic and ocular characteristics of the convert and nonconvert eyes are shown in Table 1. Convert and nonconvert eyes were similar in age, race, gender, years of follow-up, baseline IOP, IOP-lowering treatment before baseline, and IOP lowering treatment during follow-up (all comparisons, $P \geq 0.07$).

The number of convert and nonconvert eyes classified as abnormal at baseline by each classifier is shown in Table 2. All classification techniques except SVM fwd 95 were significant predictors of visual field conversion at the 0.05 probability level (Table 2). Univariate hazards ratios for the association between abnormal classification at baseline and visual field conversion for significant predictors ranged from 1.32 ($P = 0.05$) for HRTC to 2.0 ($P < 0.001$) for stereophotograph assessment. When hazards ratios for SVM classification were compared with those for other classification techniques with similar specificities (95% and 85%) in healthy eyes (see the Methods section), hazards ratios for SVM at 95% specificity (1.29, [$P = 0.07$] for SVM fwd 95 and 1.46, [$P = 0.009$] for SVM back 95) were somewhat lower than for MRA (1.54, $P = 0.003$)

TABLE 2. Comparison of Baseline Values and Cox Proportional Hazards Ratios for Univariate Associations Between Classification Techniques and Conversion to Two Consecutive Abnormal Visual Field Results

Classifier	Convert Eyes Abnormal at Baseline (n = 51)	Nonconvert Eyes Abnormal at Baseline (n = 175)	P (χ^2)	Hazards Ratio (95% CI)	P (HR)
HRTC	24 (47%)	63 (36%)	0.160	1.32 (1.00–1.75)	0.050
MRA	23 (45%)	41 (23%)	0.003	1.54 (1.16–2.02)	0.003
SVM fwd 85	43 (84%)	114 (65%)	0.080	1.85 (1.29–2.79)	<0.001
SVM fwd 95	29 (57%)	85 (49%)	0.300	1.29 (0.97–1.71)	0.070
SVM back 85	43 (84%)	122 (70%)	0.030	1.69 (1.19–2.56)	0.003
SVM back 95	33 (65%)	89 (51%)	0.080	1.46 (1.10–1.97)	0.009
Stereophotographs	33 (69%)*	73 (42%)†	0.001	2.0 (1.48–2.76)	<0.001

* n = 48, † n = 173. Five eyes had poor-quality stereophotographs that were omitted from the analysis.

(the parameter with a similar specificity in healthy eyes). The hazards ratios for SVM at 85% specificity (1.69, [P = 0.003] for SVM back 85 and 1.85 [P < 0.001] for SVM fwd 85) were somewhat higher than for HRTC and lower than for stereophotograph assessment— both parameters with similar specificity in healthy eyes. The univariate hazards ratios for age at baseline HRT image (per 1 year older), closest SAP PSD (per 1 dB higher), and closest IOP (per 1 mm Hg higher) were 1.04 (CI = 1.01–1.07; P = 0.003), 5.13 (CI = 2.33–11.26; P < 0.001), and 0.97 (CI = 0.93–1.03; P = 0.3), respectively.

When each of the six HRT predictors and stereophotography results were included individually in multivariate proportional hazards models with age, SAP PSD, and IOP; MRA (hazards ratio [HR] = 1.47), SVM fwd 85 (HR = 1.58), SVM back 85 (HR = 1.51), and stereophotograph classification (HR = 1.90) were significant predictors of visual field abnormalities (all P < 0.02). SVM fwd 95 (H.R = 1.19, P = 0.22) was not a significant independent predictor of repeatable glaucomatous visual field abnormality, whereas HRTC (HR = 1.30, P = 0.06) and SVM back 95 (HR = 1.30, P = 0.07) approached significance (Table 3). Figure 1 shows Kaplan-Meier survival curves for the classification techniques that were significant independent predictors of repeatable glaucomatous visual field abnormalities (MRA, SVM fwd 85, SVM back 85, and stereophotograph assessment). These curves show the cumulative survival (nonconvert) rate during follow-up for eyes classified as abnormal at baseline compared with those classified as normal at baseline.

Baseline age was a significant predictor in multivariate models with HRTC, SVM fwd 95, and SVM back 95 (all HR = 1.03 per 1-year increase in age, all P < 0.05); baseline SAP PSD was a significant predictor in all multivariate models (HR = 3.3–4.7 per 1-dB increase in PSD, all P < 0.004); and baseline IOP was never a significant predictor in multivariate models (all P > 0.25).

It is possible that including SAP PSD as a predictor in the multivariate models is an overadjustment, because PSD (at time of conversion) was part of our definition of visual field abnormality. To address this, we reanalyzed the data omitting PSD from the models (copredictors were age and IOP). Significant predictors were MRA (HR = 1.48, P = 0.007), SVM fwd 85 (HR = 1.71, P = 0.002), SVM back 85 (HR = 1.59, P = 0.008), SVM back 95 (HR = 1.38, P = 0.03) and stereophotograph assessment (HR = 1.99, P < 0.001) (Table 3). In addition, age was a significant predictor in all models (HR = 1.03–1.04 per 1-year increase in age, all P < 0.03), and IOP was never a significant predictor (all P > 0.50).

Convert 3 Criterion

For the convert 3 criterion, 37 (16%) of 226 eyes had three consecutive abnormal SAP results, and 189 (84%) of 226 eyes did not. Demographic and ocular characteristics of the convert and nonconvert eyes are shown in Table 1 and are similar to those for the convert 2 criterion.

The number of convert and nonconvert eyes classified as abnormal at baseline by each classifier for the convert 3 criterion are shown in Table 4. Significant univariate predictors of visual field conversion at $P \leq 0.05$ were SVM fwd 85 (HR = 1.73, P = 0.006), SVM back 85 (HR = 1.82, P = 0.004), and stereophotograph assessment (HR = 1.77, P = 0.001; Table 3). Hazards ratios for SVM at 95% specificity (1.13, [P = 0.44] for

TABLE 3. Multivariate Cox Proportional Hazards Ratios between Classification Techniques and Conversion to Two Consecutive Abnormal Visual Field Results

Classifier	Hazards Ratio (95% CI)	P (Likelihood Ratio χ^2)	Hazards Ratio without SAP PSD	P
HRTC	1.30 (0.99–1.73)	0.060	1.31 (0.99–1.73)	0.060
MRA	1.47 (1.10–1.95)	0.009	1.48 (1.11–1.96)	0.007
SVM fwd 85	1.58 (1.10–2.42)	0.010	1.71 (1.19–2.60)	0.002
SVM fwd 95	1.19 (0.90–1.58)	0.220	1.21 (0.92–1.62)	0.170
SVM back 85	1.51 (1.06–2.29)	0.020	1.59 (1.12–2.42)	0.008
SVM back 95	1.30 (0.97–1.76)	0.070	1.38 (1.04–1.87)	0.030
Stereophotographs*	1.90 (1.37–2.69)	<0.001	1.99 (1.44–2.81)	<0.001

Data were obtained after controlling for age (y), SAP PSD result (dB), and IOP (mm Hg) (column 2) and after controlling for age and IOP only (column 4), in the multivariate model.

* n = 48 converts and n = 173 nonconverts. Five eyes had poor-quality stereophotographs that were omitted from the analysis.

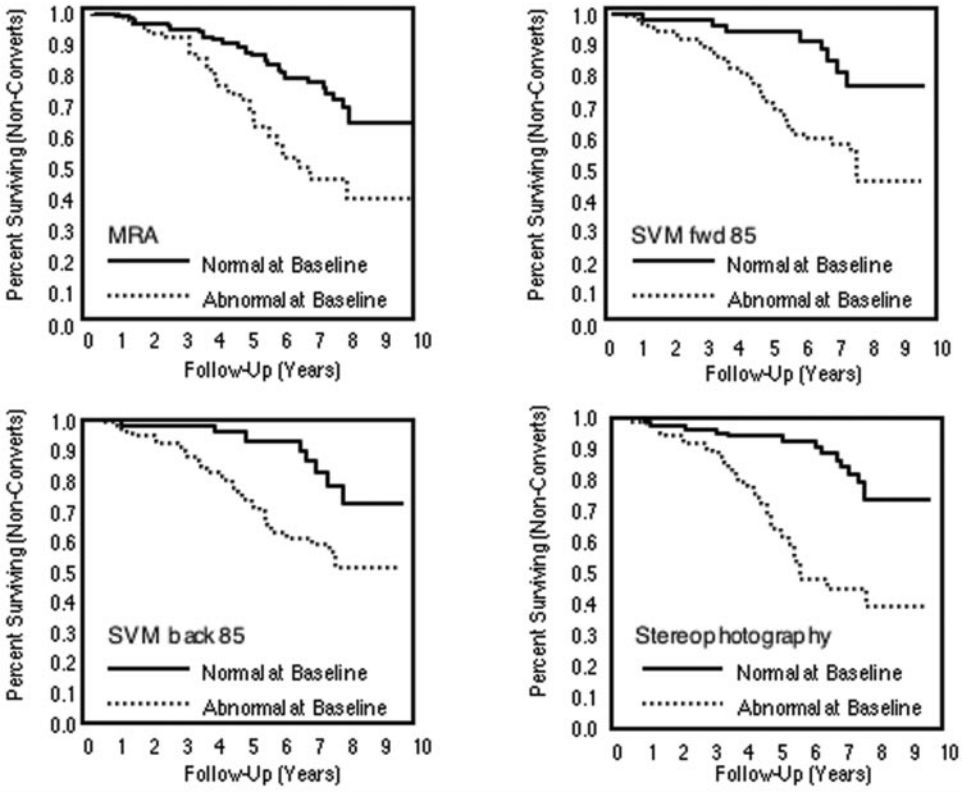

FIGURE 1. Kaplan-Meier cumulative survival curves for the four classification techniques that were significant independent predictors of the development of two consecutive abnormal visual field rest results (MRA, SVM 85 fwd, SVM 85 back, and stereophotography). Other parameters in the multivariate proportional hazards regression models were age, SAP PSD, and IOP.

SVM fwd 95 and 1.28, [$P = 0.14$] for SVM back 95) were generally similar to MRA (1.27, $P = 0.17$), the parameter with a similar specificity in healthy eyes. The hazards ratios for SVM at 85% specificity (1.73, [$P = 0.006$] for SVM fwd 85 and 1.82, [$P = 0.004$] for SVM back 85) were higher than for HRTC (HR = 0.98, $P = 0.90$) and similar to stereophotograph assessment. The univariate hazards ratios for age at baseline HRT image (per 1 year older), closest SAP PSD (per 0.1 dB higher), and closest IOP (per 1 mm Hg higher) were 1.03 (CI = 1.00–1.06; $P = 0.09$), 2.96 (CI = 1.15–7.60; $P = 0.02$), and 0.98 (CI = 0.92–1.04; $P = 0.49$), respectively.

When each of the six HRT predictors and stereophotography results were included individually in multivariate proportional hazards models with age, SAP PSD, and IOP, SVM fwd 85 (HR = 1.57, $P = 0.03$), SVM back 85 (HR = 1.69, $P = 0.02$), and stereophotograph assessment (HR = 1.70, $P = 0.004$) were the only significant independent predictors of repeatable glaucomatous visual field abnormalities (all others $P > 0.32$). These results are shown in Table 5, and Kaplan-Meier survival curves for these classification techniques are shown in Figure 2.

For the convert 3 criterion, SAP PSD was a borderline independent predictor ($0.10 > P > 0.05$) in multivariate models with HRTC, MRA, SVM 85, SVM 95, SVM fwd 95, and SVM back 95 (HR = 2.4–2.7 per 1 dB increase in PSD). Neither age nor IOP was a significant predictor in any multivariate models based on the convert 3 criterion (all $P > 0.25$).

When classifiers were assessed in multivariate models excluding PSD (copredictors age and IOP only), significant predictors were SVM fwd 85 (HR = 1.65, $P = 0.02$), SVM back 85 (HR = 1.75, $P = 0.01$), SVM back 85 (HR = 1.75, $P = 0.009$), and stereophotograph assessment (HR = 1.75, $P = 0.002$) (Table 5). In addition, age was a significant predictor in all models (HR = 1.03–1.04 per 1-year increase

TABLE 4. Comparison of Baseline Values and Cox Proportional Hazards Ratios for Univariate Associations between Classification Techniques and Conversion to Three Consecutive Abnormal Visual Field Results

Classifier	Convert Eyes Abnormal at Baseline ($n = 37$)	Nonconvert Eyes Abnormal at Baseline ($n = 189$)	P (χ^2)	Hazards Ratio (95% CI)	P (HR)
HRTC	13 (35%)	74 (39%)	0.64	0.98 (0.67–1.36)	0.900
MRA	14 (38%)	50 (26%)	0.16	1.27 (0.90–1.76)	0.170
SVM fwd 85	31 (84%)	126 (67%)	0.03	1.73 (1.16–2.82)	0.006
SVM fwd 95	19 (51%)	95 (50%)	0.90	1.13 (0.82–1.57)	0.440
SVM back 85	32 (86%)	133 (70%)	0.03	1.82 (1.19–3.12)	0.004
SVM back 95	22 (59%)	100 (53%)	0.46	1.28 (0.92–1.79)	0.140
Stereophotographs	23 (66%)*	83 (45%)†	0.02	1.77 (1.26–2.56)	0.001

* $n = 35$, † $n = 186$. Five eyes had poor-quality stereophotographs: that were omitted from the analysis.

TABLE 5. Multivariate Cox Proportional Hazards Ratios between Classification Techniques and Conversion to Three Consecutive Abnormal Visual Field Results

Classifier	Hazards Ratio (95% CI)	*P* (Likelihood Ratio χ^2)	Hazards Ratio without SAP PSD	*P*
HRTC	0.93 (0.65–1.30)	0.66	0.94 (0.66–1.32)	0.73
MRA	1.19 (0.83–1.66)	0.33	1.21 (0.85–1.69)	0.28
SVM fwd 85	1.57 (1.03–2.59)	0.03	1.65 (1.09–2,71)	0.02
SVM fwd 95	1.05 (0.75–1.46)	0.77	1.08 (0.77–1.51)	0.64
SVM back 85	1.69 (1.09–2.90)	0.02	1.75 (1.14–3.00)	0.009
SVM back 95	1.16 (0.83–1.64)	0.38	1.22 (0.87–1.72)	0.24
Stereophotographs*	1.70 (1.18–2.51)	0.004	1.75 (1.22–2.59)	0.002

Data were obtained after controlling for age (y), SAP PSD result (dB), and IOP (mm Hg) (column 2) and after controlling for age and IOP only (column 4), in the multivariate model.

* $n = 35$ converts and $n = 186$ non-converts. Five eyes had poor-quality stereophotographs that were omitted from the analysis.

in age, all $P < 0.03$), and IOP was never a significant predictor (all $P > 0.50$).

DISCUSSION

Our results suggest that both baseline structural and functional measurements predict the development of repeatable visual field damage. Moreover, in our population of glaucoma-suspect eyes, disc measurements were predictive of the development of visual field abnormalities, even after adjusting for functional measures. Thus, the degree of baseline disc damage measured by HRT and observed by stereophotograph assessment was predictive of the development of abnormal visual fields. Although SVM classifiers trained on HRT data were somewhat more predictive of the development of visual field abnormalities than the HRTC, they were not more predictive than stereophotograph evaluation by experts or the HRT MRA.

Univariate analyses indicated that glaucoma-suspect eyes with baseline abnormal HRTC, abnormal MRA result, abnormal Gaussian SVM classification, or glaucomatous-appearing optic discs by stereophotography, have a 32% (HRTC) to 100% (stereophotography) increased risk of having of two consecutive abnormal visual field results (conversion criterion 2) compared with eyes that are normal at baseline. The risk for three consecutive abnormal field results (criterion 3) was smaller for baseline abnormality with these classification techniques, ranging from −2% (nonsignificant, HRTC) to 77% (stereophotography). Lower hazards ratios for three consecutive abnormal results are probably due to the classification of eyes with two consecutive abnormal field re-

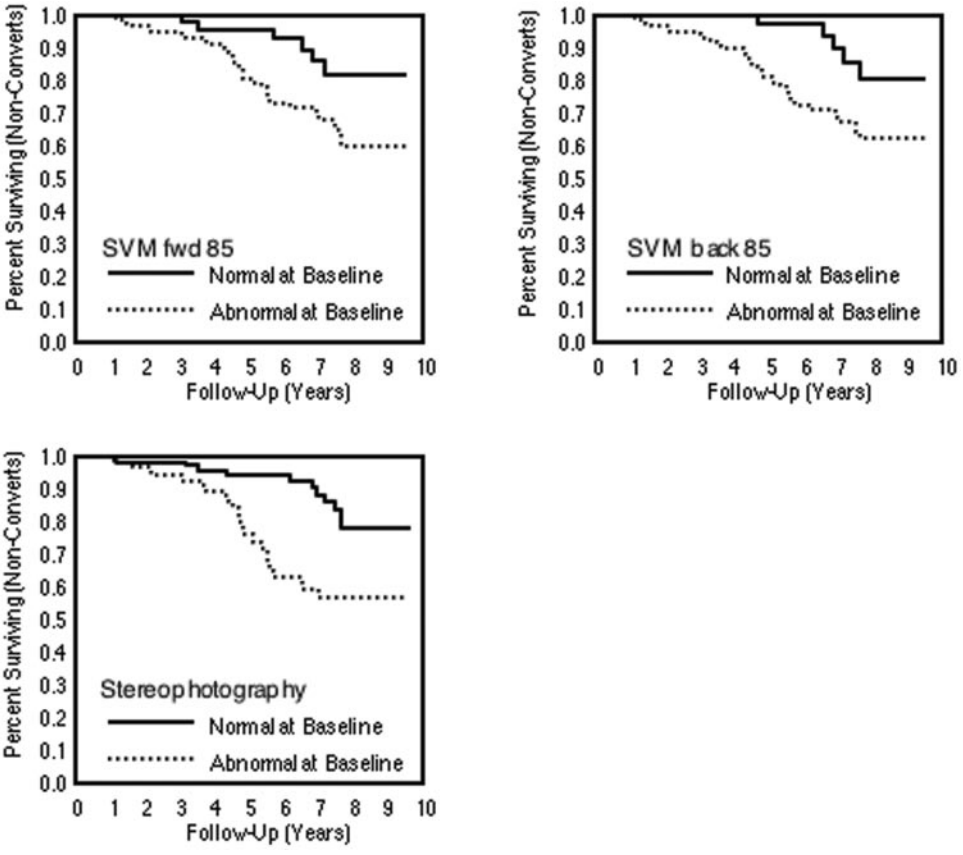

FIGURE 2. Kaplan-Meier cumulative survival curves for the three classification techniques that were significant independent predictors of the development of three consecutive abnormal visual field rest results (SVM 85 fwd, SVM 85 back, and stereophotography). Other parameters in the multivariate proportional hazards regression models were age, SAP PSD, and IOP.

sults as nonconvert eyes. It is possible that eyes with two consecutive abnormal results are more likely to have a third abnormal result, than eyes with one abnormal result are likely to have a second.

When other predictor variables (age, SAP PSD, and IOP) were included in multivariate models, fewer HRT classifiers were significant independent predictors for both conversion criteria 2 and 3. Only SVM fwd 85, SVM back 85, and stereophotograph assessment, remained significant univariate and independent multivariate predictors for both convert criteria. In addition, for the criterion 2 condition, SAP PSD was a significant predictor when included in multivariate models with each of the optic disc classification techniques.

It is interesting to note that baseline visual field indices that fell within normal limits (all baseline visual fields were normal by PSD), but may have had values approaching the normal–abnormal cutoff, were predictors of visual field abnormalities in conjunction with structural measures. Due to the high variability of visual field tests, however, it may be unlikely that small differences in PSD are clinically useful predictors of visual field damage. Alternatively, perhaps there is information in the visual field that is not fully used by currently available analysis techniques.

We also found that abnormal appearing optic discs, classified by expert stereophotograph graders, were significant predictors of visual field abnormalities in univariate and multivariate models, as previously reported.[25] Although disc photograph assessment was as strong or stronger a predictor of visual field conversion (based on hazard ratio magnitude) than HRT assessment, HRT has several advantages over stereophotograph assessment. The HRT provides immediate objective results, and its use requires little expertise, whereas optic disc photograph assessment requires processing time and subjective evaluation by trained experts. HRT also is likely less affected by ocular media, lens opacities, and poor pupil dilation.

SVMs were significant predictors of visual field conversion in the present study, with univariate and multivariate hazards ratios reported for classification using SVM similar to those using MRA, but somewhat smaller than those using stereophotograph assessment. Stereophotograph assessment may have benefited from the considerable experience of the expert graders. In addition, SVM cutoffs used in the present study were derived from SVMs originally trained to discriminate between healthy eyes and those with early to moderate glaucomatous field loss (average MD = −5.74 dB)[20] probably not optimizing their sensitivity for classifying preperimetric suspect eyes.

Results from the present study suggest that Gaussian SVM classifiers perform as well as, or better than, other commercially available classification techniques provided by HRT software at "predicting" which glaucoma-suspect eyes will later undergo development of repeatable visual field abnormalities measured using SAP. In a related study, Sample et al.[26] showed that SVMs, trained to discriminate between healthy eyes and those with glaucomatous-appearing optic discs by stereophotograph assessment, identified abnormality in longitudinally observed visual field converts at approximately the same time as standard methods. In addition, two previous studies compared SVMs to commercially available classification techniques for discriminating between healthy and glaucomatous eyes. Bowd et al.[12] showed that SVM trained on global and regional HRT parameters improves on previously reported HRT parameter-based linear discriminant functions for discriminating between healthy eyes and those with repeatable (two consecutive) visual field abnormalities. ROC curve areas were significantly larger and sensitivities at fixed specificities were higher for SVM classification than for HRTC, and several other published HRT parameter-based linear discriminant functions. Goldbaum et al.[19] used SVM to classify SAP 24-2 visual fields, based on absolute threshold at all 52 test points, from healthy eyes and those with glaucomatous appearing optic discs. In that study, ROC curve areas, sensitivities, and specificities for SVM classification were found as good as, or better than, those for STATPAC (Carl Zeiss Meditec) global indices (MD, PSD, CPSD, GHT) and human experts. These results suggest that SVMs are robust techniques for both structural and functional classification problems in glaucoma.

Other studies have shown change in HRT-measured disc topography occurring at the same time as, or before, visual field change. Kamal et al.[27] showed change in HRT parameters in ocular hypertensive eyes that converted to glaucomatous visual fields. They determined the parameter variability between two images in healthy eyes and compared this variability to that shown between two images from convert eyes obtained at preconvert baseline and at the time of the third repeatable abnormal visual field result. If the change in measurements between the two images was greater than the variability in healthy eyes, the eye was classified as "getting worse." Thirteen of 21 convert eyes showed change in this direction. In another study, Tan and Hitchings[10] determined the limits of variability in 30° neuroretinal rim sectors of individual study eyes, using an eye-specific reference plane, and classified eyes as "progressing" if measurements in a single sector changed more than the limit of variability, in two of three consecutive tests. This technique had sensitivity of 85% for detecting 20 eyes with no visual field defects that later showed visual field defects later on SAP. Specificity was 95% in healthy eyes followed over time. Finally, Chauhan et al.[11] detected retinal height (disc and parapapillary area) changes (three repeatable outside the limits of test–retest variability) in 69% of 77 patients with glaucoma observed longitudinally for approximately 5 years. Twenty-seven percent of the same eyes showed SAP changes based on the commercially available Glaucoma Change Probability Analysis requiring worsening of four identical repeatable points in two of three follow-up examinations. However, in patients showing disc and field changes, the disc changes did not significantly precede the field changes. Further, increased change with HRT compared with SAP may have been a result of the incompatible criteria used to define change across instruments. Agreement between HRT and expert stereophotograph evaluation for detecting progression was good in the subset of eyes that had both tests (13/16, 81%).

In the present study, we found a significant percentage of converts with normal baseline classification results and a significant percentage of nonconverts with abnormal baseline results, when using both HRT MRA and SVMs. Therefore, glaucomatous visual fields developed in some eyes with normal baseline disc measurements but not in some eyes with abnormal baseline disc measurements. These results suggest that the classification techniques used are not optimized for identifying glaucoma converts based on preconversion measurements. Future research should concentrate on improving or developing new classification techniques for this purpose.

It is important to recognize that, although baseline optic disc abnormalities were present in many of the eyes evaluated, it does not mean that these abnormalities always precede the development of glaucomatous visual fields. If eyes with normal discs at baseline were selected for study, instead of suspect eyes including those with abnormal-appearing optic discs, it is likely that in some cases development of visual field abnormalities would precede disc abnormalities.[14] Therefore, our results cannot be generalized to a random sample of eyes, rather a sample that is more representative of one observed in a glaucoma clinic.

In conclusion, both stereophotography and HRT disc abnormalities can precede and predict repeatable visual field abnormalities measured using SAP, the current clinical stan-

dard for diagnosis, although the strength of these predictors is not absolute. These results support the importance of optic disc examination for early glaucoma diagnosis.

References

1. Sommer A, Miller NR, Pollack I, Maumenee AE, George T. The nerve fiber layer in the diagnosis of glaucoma. *Arch Ophthalmol*. 1977;95:2149–2156.
2. Quigley HA, Addicks EM, Green WR. Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy. *Arch Ophthalmol*. 1982;100:135–146.
3. Zeyen TG, Caprioli J. Progression of disc and field damage in early glaucoma. *Arch Ophthalmol*. 1993;111:62–65.
4. Bowd C, Zangwill LM, Blumenthal EZ, et al. Imaging of the optic disc and retinal nerve fiber layer: the effects of age, optic disc area, refractive error, and gender. *J Opt Soc Am A Opt Image Sci Vis*. 2002;19:197–207.
5. Garway-Heath DF, Wollstein G, Hitchings RA. Aging changes of the optic nerve head in relation to open angle glaucoma. *Br J Ophthalmol*. 1997;81:840–845.
6. Gundersen KG, Heijl A, Bengtsson B. Age, gender, IOP, refraction and optic disc topography in normal eyes: a cross-sectional study using raster and scanning laser tomography. *Acta Ophthalmol Scand*. 1998;76:170–175.
7. Nakamura H, Maeda T, Suzuki Y, Inoue Y. Scanning laser tomography to evaluate optic discs of normal eyes. *Jpn J Ophthalmol*. 1999;43:410–414.
8. Saruhan A, Orgul S, Kocak I, Prunte C, Flammer J. Descriptive information of topographic parameters computed at the optic nerve head with the Heidelberg retina tomograph. *J Glaucoma*. 1998;7:420–429.
9. Kamal DS, Viswanathan AC, Garway-Heath DF, Hitchings RA, Poinoosawmy D, Bunce C. Detection of optic disc change with the Heidelberg retina tomograph before confirmed visual field change in ocular hypertensives converting to early glaucoma. *Br J Ophthalmol*. 1999;83:290–294.
10. Tan JC, Hitchings RA. Approach for identifying glaucomatous optic nerve progression by scanning laser tomography. *Invest Ophthalmol Vis Sci*. 2003;44:2621–2626.
11. Chauhan BC, McCormick TA, Nicolela MT, LeBlanc RP. Optic disc and visual field changes in a prospective longitudinal study of patients with glaucoma: comparison of scanning laser tomography with conventional perimetry and optic disc photography. *Arch Ophthalmol*. 2001;119:1492–1499.
12. Bowd C, Chan K, Zangwill LM, et al. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Invest Ophthalmol Vis Sci*. 2002;43:3444–3454.
13. Zangwill LM, Bowd C, Berry CC, et al. Discriminating between normal and glaucomatous eyes using the Heidelberg Retina Tomograph, GDx Nerve Fiber Analyzer, and Optical Coherence Tomograph. *Arch Ophthalmol*. 2001;119:985–993.
14. Gordon MO, Beiser JA, Brandt JD, et al. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. *Arch Ophthalmol* 2002;120:714–720; discussion 829–830.
15. Weinreb RN, Lusky M, Bartsch DU, Morsman D. Effect of repetitive imaging on topographic measurements of the optic nerve head. *Arch Ophthalmol*. 1993;111:636–638.
16. Iester M, Mikelberg FS, Drance SM. The effect of optic disc size on diagnostic precision with the Heidelberg retina tomograph. *Ophthalmology*. 1997;104:545–548.
17. Wollstein G, Garway-Heath DF, Hitchings RA. Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology*. 1998;105:1557–1563.
18. Bishop CM. *Neural Networks for Pattern Recognition.* Oxford, UK: Clarendon Press; 1995.
19. Goldbaum MH, Sample PA, Chan K, et al. Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. *Invest Ophthalmol Vis Sci*. 2002;43:162–169.
20. Zangwill LM, Chan K, Bowd C, et al. Heidelberg retina tomograph measurements of the optic disc and parapapillary retina for detecting glaucoma analyzed by machine classifiers. *Invest Ophthalmol Vis Sci.* In press.
21. Zangwill LM, Bowd C, Weinreb RN. Evaluating the optic disc and retinal nerve fiber layer in glaucoma II: optical image analysis. *Semin Ophthalmol*. 2000;15:206–220.
22. Park KH, Kim DM, Youn DH. Short-term change of optic nerve head topography after trabeculectomy in adult glaucoma patients as measured by Heidelberg retina tomograph. *Kor J Ophthalmol*. 1997;11:1–6.
23. Topouzis F, Peng F, Kotas-Neumann R, et al. Longitudinal changes in optic disc topography of adult patients after trabeculectomy. *Ophthalmology*. 1999;106:1147–1151.
24. Bowd C, Weinreb RN, Lee B, Emdadi A, Zangwill LM. Optic disk topography after medical treatment to reduce intraocular pressure. *Am J Ophthalmol*. 2000;130:280–286.
25. Quigley HA. Early detection of glaucomatous damage. II. Changes in the appearance of the optic disk. *Surv Ophthalmol* 1985;30:111, 117–126.
26. Sample PA, Goldbaum MH, Chan K, et al. Using machine learning classifiers to identify glaucomatous change earlier in standard visual fields. *Invest Ophthalmol Vis Sci*. 2002;43:2660–2665.
27. Kamal DS, Garway-Heath DF, Hitchings RA, Fitzke FW. Use of sequential Heidelberg retina tomograph images to identify changes at the optic disc in ocular hypertensive patients at risk of developing glaucoma. *Br J Ophthalmol*. 2000;84:993–998.